16. The method of claim **15**, wherein the particular set of software features associated with the search query is a subset of the software features of the software project, and further comprising:

generating a modified feature vector based on the feature mask;

determining a plurality of vector distances based on the modified feature vector and the subset of feature vectors resulting from the search process, each vector distance among the plurality of vector distances being determined between the modified feature vector and a respective feature vector among the subset of feature vectors; and

determining that a closest match to the software project is whichever software project among the plurality of software projects has a smallest vector distance among the plurality of vector distances.

17. The method of claim **15**, further comprising obtaining the descriptive information from one or more websites.

18. The method of claim **15**, further comprising obtaining the descriptive information about the software project from a configuration file or a readme file provided with the software project.

19. The method of claim **15**, further comprising determining the software features of the software project by:

determining a plurality of counts corresponding to a plurality of textual terms in the descriptive information, each respective count among the plurality of counts indicating how many times a respective textual term is present in the descriptive information;

determining a subset of textual terms from among the plurality of textual terms having respective counts above a predefined threshold; and

designating the subset of textual terms as the software features for the software project.

20. A non-transitory computer-readable medium comprising program code that is executable by a processor for causing the processor to:

analyze descriptive information about a software project to determine software features of the software project, the software features being functional characteristics of the software project;

generate a feature vector for the software project based on the software features of the software project, the feature vector being a data structure of elements in which each element is a numerical value indicating whether a particular software feature corresponding to the element is among the software features of the software project as determined from the descriptive information; and

store the feature vector in a database having a plurality of feature vectors for a plurality of software projects, the plurality of feature vectors being searchable in response to search queries.

* * * * *